PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C07K 14/315, A61K 39/09, C07K 16/12

A1

(11) International Publication Number: WO 00/37496

(43) International Publication Date: 29 June 2000 (29.06.00)

(21) International Application Number: PCT/SE99/02448

(22) International Filing Date: 21 December 1999 (21.12.99)

(30) Priority Data:
9804491-0 22 December 1998 (22.12.98) SE

(71)(72) Applicants and Inventors: GUSS, Bengt [SE/SE]; Dag Hammarskjölds väg 238 B, S-756 52 Uppsala (SE). LINDMARK, Hans [SE/SE]; Tiundagatan 61 2 tr, S-752 39 Uppsala (SE). JACOBSSON, Karin [SE/SE]; Stabby Allé 9 C, S-752 29 Uppsala (SE). FRYKBERG, Lars [SE/SE]; Stabby Allé 7 C, S-752 29 Uppsala (SE).

(74) Agent: AGVALD-GLAS, Gunilla; AB Stockholms Patentbyrå, Zacco & Bruhn, P.O. Box 23101, S-104 35 Stockholm (SE).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: NOVEL FIBRONECTIN-BINDING PROTEIN

(57) Abstract

The present invention is concerned with a novel fibronectin-binding protein of *Streptococcus equi*, to a DNA fragment encoding this protein, to host cells and vectors containing said DNA fragment and to methods to produce said protein based on recombinant DNA technology. The invention is also related to use of said protein in the preparation of a vaccine, to a vaccine containing said protein, to antibodies specific for said protein and to polyvalent antisera containing such antibodies.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Novel fibronectin-binding protein

The present invention is generally related to a novel protein, methods to produce said protein and use thereof, e. g. for immunization purposes.

More specifically, the present invention is related to a novel fibronectin-binding protein derived from a bacterium belonging to the genus *Streptococcus,* to a DNA sequence encoding said protein, recombinant DNA methods for the production of said protein, and use of said protein per se or a fragment thereof as an immunogenic protein or antigenic polypeptide or peptide, e. g. for use as an active component in a vaccine, or to produce antisera.

Streptococcal infections in horses are mainly caused by the species *Streptococcus equi*, which is classified as a Lancefield Group C *Streptococcus* and comprises two subspecies designated *equi* and *zooepidemicus,* respectively. *Streptococcus equi* subsp. *equi* which is virtually confined to horses is the causative agent of strangles, a world-wide distributed and serious disease of the equine upper respiratory tract. Since strangles is a highly contagious disease, not only infected animals but also all other members of an afflicted stud must be isolated for as long as up to three months.

S. equi subsp. *zooepidemicus*, is considered as an opportunistic commensal often occurring in the upper respiratory tract of healthy horses. However, after stress or virus infection, it can cause a secondary infection, which results in strangles-like symptoms. Moreover, subsp. *zooepidemicus* infects not only horses but also a wide range of other animals, like pigs, dogs, cats, and cows. Even human cases with infection due to subsp. *zooepidemicus* have been reported. This subspecies has been implicated as the primary pathogen in conditions such as endometritis, cervicitis, abortion, mastitis, pneumonia, abscesses and joint infections.

Although it is possible to treat and cure these streptococcal infections with antibiotics, such as penicillin, tetracycline or gentamicin, an effective prophylactic agent, that could prevent outbursts of such infections and obviate or reduce the risk for development of resistant strains associated with antibiotic treatment, would be appreciated.

However, although many attempts have been made to develop prophylactic agents such as vaccines against *S. equi,* at the present time no efficient vaccines or immunizing preparations are available, neither for the subspecies *equi* nor for the subspecies *zooepidemicus*.

Existing vaccines against strangles are based on inactivated, e. g. heat- killed, or attenuated strains of *S. equi* subsp. *equi* or acid extracts/mutanolysin enriched in M-protein(s), i. e. immunogenic protein(s) produced by *S. equi.* A vaccine against *S. equi* subsp. *zooepidemicus* based on an M-like protein is disclosed in US-A-5 583 014.

Since the previously developed vaccines or immunizing preparations are hampered by side-effects and, moreover, provide insufficient protection, there is a need for efficient prophylactic agents, such as vaccines, that protect against *S. equi* infections and/or prevent spread thereof.

It is well known that attachment to eukaryotic cell surfaces is an essential step in the establishment of infection and colonization by bacterial pathogens. Accordingly, streptococcal surface proteins, that interact with and/or bind to different components of the Extracellular Matrix (ECM) or plasma proteins of the host cell, are potential candidates for use as active component(s) for immunizing purposes.

This is illustrated by the vaccines based on M-like proteins mentioned above or disclosed in the literature, i. a. in WO 98/0561. The binding of fibrinogen and complement factor H to M-proteins is assumed to be important for the ability of streptococci to resist phagocytosis by polymorphonuclear leucocytes.

Another mechanism used by streptococci for attachment to host cells involves binding to the ECM component fibronectin (Fn) (3, 4). Binding between Fn-binding bacterial cell-surface proteins and immobilized Fn promotes internalization of streptococci by epithelial cells (1, 7, 11). Fibronectin is a dimeric glycoprotein found both in plasma and in a fibrillar form in the extracellular matrix. The main function of Fn is to mediate substrate adhesion of eukaryotic cells, which involves the binding of specific cell-surface receptors to certain domains of the Fn molecule (5). Furthermore, it also interacts with several other macromolecules, such as DNA, heparin, fibrin, and collagen (5).

Accordingly, several Fn-binding proteins from different streptococcal species have been cloned and sequenced previously. From *S. equi*, one Fn-binding protein has been cloned and characterized earlier, which is a Fn-binding cell-surface protein of subsp. *zooepidemicus*, that has been designated FNZ (9).

Recently, a novel gene encoding a Fn-binding protein has been cloned from *S. equi* subsp. *equi*. The encoded protein, is clearly distinguishable from the previously isolated streptococcal Fn-binding proteins inclusive of the FNZ protein.

The present invention is based on this novel protein originally derived from *S. equi* subsp. *equi* and its potential use for immunization purposes.

Generally, the present invention is directed to a protein having an amino acid sequence encoded by a nucleic acid sequence or gene, that forms a portion of the genome of *S. equi* subsp. *equi*, and which protein binds specifically to mammalian fibronectin.

The present invention is also directed to an isolated protein, specifically binding to fibronectin, such as mammalian, and specifically equine, fibronectin, and having an amino acid sequence as shown in SEQ. ID. NO. 1.

Moreover, the present invention is generally concerned with analogs or fragments of the present protein having fibronectin-binding properties. For instance, a suitable fragment is comprised of the sequence of the present protein, that lacks the N-terminal signal sequence of the preprotein. A further suitable fragment of the present protein lacks a portion of said amino acid sequence, said portion comprising an amino acid sequence binding to a collagen-binding domain of fibronectin.

The present invention is also concerned with methods to produce the present protein, analogs or fragments thereof, which methods are based on DNA technology; and with nucleic acid sequences, and more specifically DNA sequences or fragments, intended for use in such methods, as well as use of said protein, analogs or fragments thereof for therapeutic purposes, such as immunizing purposes.

The novel protein has been termed SFS, and, accordingly, the corresponding gene is designated *sfs*. For the purpose of convenience, these terms are frequently used in the description.

In the following, the present invention is disclosed more in detail with reference to the drawings, where

Fig. 1(A) shows a map of a clone designated pSFS62 with the gene *sfs* indicated.

Fig. 1 (B) shows a schematic presentation of protein SFS with the functional domains indicated. The bars correspond to the amino acid sequences of phagemid clones isolated by panning against Fn (S1-S4). Figures refer to the amino acid positions in protein SFS as shown in SEQ. ID. NO. 1 and the figures within brackets indicate the number of identical clones, that were isolated and sequenced.

Fig. 2 shows the results of Southern blot analysis of chromosomal DNA from ten streptococcal isolates. The DNA was digested by *Apa*I and separated by pulsed-field gel electrophoresis in duplicate. The radioactively labeled probe used corresponds to the gene *sfs*. Lanes: 1, subsp. *zooepidemicus* ZV; 2, *S. dysgalactiae* S2; 3, *S. equisimilis* 172; 4, subsp. *equi* Bd 3221; 5, subsp. *equi* Bd 995; 6, subsp. *zooepidemicus* DSM 20727^T; 7, subsp.

zooepidemicus ATCC 53698; 8, subsp. *equi* CCUG 11664; 9, subsp. *equi* NCTC 9682^T; 10, *S. pyogenes* AW43. Molecular weight marker (concatamers of lamda) is indicated to the left.

Fig. 3 shows the results from inhibition assays related to Fn-binding. Cells of subsp. *zooepidemicus* ZV, subsp. *zooepidemicus* DSM 20727, subsp. *equi* Bd3221, and subsp. *equi* 640 were incubated with iodine-labeled Fn (hatched bars) and with a mixture of iodine-labeled Fn and protein SFS-E (striped bars). The bars represent means of duplicates and the standard deviation is indicated.

Fig. 4 shows the results from inhibition assays related to inhibition of binding between collagen and Fn with protein SFS. Collagen type I coated microtiter wells were incubated with Fn and a two-fold serial dilution of SFS. Bound Fn was detected by antibodies as described in Example 3. Points represent means of duplicates and the standard deviation is indicated.

More specifically, the present invention is directed to a fibronectin-binding (hereinafter abbreviated Fn-binding) protein, which has an amino acid sequence that can be expressed from a nucleic acid coding sequence, that can be isolated from and forms a portion of the genomes of *S. equi,* for instance subsp. *equi.*

According to a suitable embodiment, the present invention is directed to an isolated protein, specifically binding to fibronectin and having an amino acid sequence as shown in SEQ. ID. NO. 1 below, or a fragment or analog thereof.

SEQ. ID. NO. 1:

```
Met Arg Lys Thr Glu Gly Arg Phe Arg Thr Trp Lys Ser Lys Lys Gln
1               5                  10                  15

Trp Leu Phe Ala Gly Ala Val Val Thr Ser Leu Leu Leu Gly Ala Ala
            20                  25                  30

Leu Val Phe Gly Gly Leu Leu Gly Ser Leu Gly Gly Ser Ser His Gln
        35                  40                  45

Ala Arg Pro Lys Glu Gln Pro Val Ser Ser Ile Gly Asp Asp Asp Lys
    50                  55                  60

Ser His Lys Ser Ser Ser Asp Gln Pro Thr Asn His Gln His Gln Ala
65                  70                  75                  80

Thr Ser Pro Ser Gln Pro Thr Ala Lys Ser Ser Gly His His Gly Asn
                85                  90                  95

Gln Pro Gln Ser Leu Ser Val Asn Ser Gln Gly Asn Ser Ser Gly Gln
            100                 105                 110

Ala Ser Glu Pro Gln Ala Ile Pro Asn Gln His His Gln Pro Gln Gly
```

115 120 125

Lys Pro Gln His Leu Asp Leu Gly Lys Asp Asn Ser Ser Pro Gln Pro
130 135 140

Gln Pro Lys Pro Gln Gly Asn Ser Pro Lys Leu Pro Glu Lys Gly Leu
145 150 155 160

Asn Gly Glu Asn Gln Lys Glu Pro Glu Gln Gly Glu Arg Gly Leu Pro
165 170 175

Gly Leu Asn Gly Glu Asn Gln Lys Glu Pro Glu Gln Gly Glu Arg Gly
180 185 190

Glu Ala Gly Pro Pro Ser Thr Pro Asn Leu Glu Gly Asn Asn Arg Lys
195 200 205

Asn Pro Leu Lys Gly Leu Asp Gly Glu Asn Lys Pro Lys Glu Asp Leu
210 215 220

Asp Gly Tyr Asn His Gly Arg Arg Asp Gly Tyr Arg Val Gly Tyr Glu
225 230 235 240

Asp Gly Tyr Gly Gly Lys Lys His Lys Gly Asp Tyr Pro Lys Arg Phe
245 250 255

Asp Glu Ser Ser Pro Lys Glu Tyr Asn Asp Tyr Ser Gln Gly Tyr Asn
260 265 270

Asp Asn Tyr Gly Asn Gly Asn Pro Asp
275 280

The present invention is also related to proteins or polypeptides having an amino acid sequence as shown in SEQ. ID. NO. 1 containing deletions or substitutions of amino acids, such as fragments and analogs of the present protein having fibronectin-binding properties, suitably conserved, or specifically designed, Fn-binding properties. One such fragment or analog is comprised of the mature protein lacking the N-terminal amino acids no. 1 to 29 inclusive. Other fragments have en amino acid sequence corresponding to a portion of the amino acid sequence as shown in SEQ. ID. NO. 1 comprising a fibronectin-binding domain or an antigenic determinant or epitope. Still other fragments have an amino acid sequence corresponding to a portion of the sequence as shown in SEQ. ID. NO. 1, wherein an amino acid sequence binding to a collagen-binding domain of fibronectin (Fn) and comprising the amino acids QGERGEAGPP, is deleted.

A further embodiment is concerned with a protein of the present invention having an amino acid composition of approximately 53 glycine residues, 39 serine residues and 38 proline residues evenly distributed in the protein and optionally 13 tyrosine residues in the C-terminal part of the protein.

Obviously, the present invention is concerned with a wild-type ptotein encoded by *S. equi,* that can be isolated and purified when recovered from said organism, as well as with a recombinant SFS protein as discussed above, said proteins having Fn-binding properties.

The present invention is also concerned with a nucleic acid sequence encoding the SFS protein or fragments or analogs thereof. Suitably, this sequence is a DNA sequence and contains an SFS coding sequence, such as the entire *sfs* gene, or a portion thereof encoding the SFS protein or a fragment or analog therof.

Accordingly, one embodiment of the present invention is related to a DNA sequence having a nucleotide sequence as shown in SEQ, ID. NO. 2 or to an equivalent thereof.

SEQ. ID. NO. 2:

```
ATGAGAAAAA CAGAAGGACG TTTTCGCACA TGGAAGTCCA AAAAACAATG GCTATTTGCC    60
GGTGCAGTGG GAGCTGCACT TGTCTTTGGA GGTTTATTAG GAAGTCTTGG TGGCTCATCC   120
CAGCAGCCAG TCAGCTCGAT TGGAGATGAC GATAAGTCGC ACAAGAGCTC ATCACCACCG   180
AAAAAGGATA ACTTGCAGCC TAAGCCTTCA GATCAGCCTA CTAATCGCCC GTCCCAGCCG   240
ACAGCAAAGA GCTCAGGTCA TCATGGGAAT CAACCTCACC AAGGAAATAG TAGTGGACAG   300
GCCTCAGAGC CTCAGGCTAT TCCTAATCAA GGGCTGAGAG GAGGTAACAG CTCTGGTTCA   360
GGTCATCACC ATCAGCCACA AGATCTAGGT AAGGATAATT CTAGCCCGCA GCCTCAACCA   420
AAGCCTCAGG GCAAAAAAGG CTTGAATGGT GAAAATCAGA AGGAACCGGA GCAAGGTGAA   480
CGAGGTTCAG GGTTGAGTGG TAATAATCAA GGCCGTCCTT CGCTTCCAGG CTTGAATGCA   540
GAGCAAGGTG AACGAGGTGA AGCCGGTCCC CCATCAACTC CGAATTTAGA TCCTTTAAAA   600
GGATTAGATG GAGAGAATAA GCCAAAGGAA GATTTAGACG GTAATGATGA ATCACCAAAA   660
CTTAAAGACG AACACCCCTA CAATCATGGT CGTCGCTATG AGGATGGATA TGGTGGCAAA   720
AAGCACAAAG GAGATTATCC TAAGCGAAAG GAATATAATG ACTATAGTCA AGGGTATAAT   780
GATAATTATG GAAATGGCGA TAGAGGTGGT AAGAGAGGAT ACGGCTATTC TTACAATCCC   840
GACTAA                                                              846
```

The present invention is also related DNA sequences having nucleotide substitutions, that do not change the encoded product or interfere with its expression. This is due to the well-known redundancy of the genetic code, i. e. more than one coding nucleotide triplet (codon) can code for or define a particular amino acid residue or a function, such as a stop codon function, etc.. Thus, such functionally equivalent sequences are also encompassed by the present invention. Such equivalents may also arise due to spontaneous mutations.

Accordingly, when used to produce the present protein, such as a protein having the amino acid sequence shown in SEQ. ID. NO. 1, with methods based on recombinant

DNA technology, the above DNA sequence may be modified to adapt to the codon frequency of the host organism used to produce the present protein, or fragments or analogs thereof.

The present invention is also concerned with a host cell comprising a DNA fragment or sequence of the present invention, e. g. a eukaryotic or, suitably, a prokaryotic host cell. A suitable prokaryotic host cell is derived from different strains of *E. coli*.

Furthermore, the present invention is concerned with a method to produce the present protein, fragments or analogs thereof comprising culturing a host cell containing the DNA sequence of the present invention, and isolating the expressed protein from the culture. Suitably this method also comprises purification of the expression product, such as affinity chromatography purification, conveniently based on use of "affinity tails".

Thus, a suitable method comprises

(a) introducing a DNA fragment of the present invention into an expression vector;

(b) introducing the said vector, which contains the said DNA fragment, into a compatible host cell;

(c) culturing the host cell provided in step (b) under conditions required for expression of the product encoded by said DNA fragment; and

(d) isolating the expressed product from the cultured host cell.

Preferably, said method further comprises a step (e), wherein the isolated product from step (d) is purified, e. g. by affinity chromatography, such as Fn-affinity chromatography, or with the use of "affinity tails" as is well-known in this field of art.

The present invention is further concerned with a vaccine comprising the present Fn-binding protein or a fragment or an analog thereof as an antigenic or immunogenic component. This vaccine is intended for use as a vaccine protecting against infection with any one of the two subspecies *equi* and *zooepidemicus* of *S. equi*.

A further embodiment of the present invention is concerned with a vaccine as defined above, that protects horses against strangles caused by *S. equi* subsp. *equi* infection. Suitably, such a vaccine could protect also against infection with subsp. *zooepidemicus*.

The vaccine of the present invention is suitably a subunit vaccine. Moreover, the vaccine may be comprised of a cocktail of antigenic and /and or immunogenic components comprising the present protein or an analog or a fragment thereof as one such component.

The present invention is also concerned with use of the present Fn-binding protein, analogs or fragments thereof in the preparation of a vaccine protecting against S. *equi*, for instance a vaccine as disclosed above.

Furthermore, the present invention is also concerned with the production of antibodies raised against the present protein , analogs or fragments thereof, with fragments of such antibodies and with the production of antisera. Antibodies and/or antisera could be produced by in vivo administration, e. g. injection, of an antigen comprising the said protein, an analog or a fragment thereof, to a host to elicit an immune response in said host, and recovering antiserum thereby produced in said host and, optionally, recovering or isolating antibodies contained in said antiserum or in other body fluids from said host. Not only polyclonal but also monoclonal antibodies could be produced in accordance with well-known methods.

EXPERIMENTAL PART

Bacterial strains, plasmids, and growth conditions. Ninety-eight clinical isolates of *S. equi* (50 isolates of subsp. *equi* strains and 48 isolates of subsp. *zooepidemicus* strains) collected from different parts of Sweden between 1982-1996 were together with the streptococcal control strains *S. dysgalactiae* S2 and *S. equisimilis* 172 obtained from the National Veterinary Institute, Uppsala. The *S. pyogenes* strain AW-43 was a kind gift from Dr G. Lindahl, Lund University. The plasmid pUC19 was used for cloning purposes and pGEX-5X-2 (Pharmacia Biotech, Uppsala, Sweden) for facilitating purification of proteins. Phagemid pG8SAET (19) was used for purification of protein SFS and for construction of the phage display library. Streptococcal strains were grown on horse blood agar plates or in Todd-Hewitt broth (Oxoid, Basingstoke, UK) supplemented with 0.5% yeast extract (THY). *E. coli* strains were cultured in Luria-Bertani (LB) medium supplemented in appropriate cases with 50 µg of ampicillin per ml.

As regards the isolates of *S. equi,* it is known that isolates of subsp. *equi* are serologically and genetically very homogeneous whereas isolates of subsp. *zooepidemicus* display a high degree of heterogeneity (2, 8, 10, 13).

Example 1. Cloning and isolation of a gene *sfs* encoding the Fn-binding protein SFS.

A. Construction of a phagemid library. A shotgun phage display library was constructed from subsp. *equi* Bd 3221 essentially as described by Jacobsson and Frykberg (6). Briefly, chromosomal DNA of this strain was isolated as described earlier (9) and subsequently purified and fragmented by sonication.

The obtained fragments were treated with T4 DNA polymerase to generate blunt ends and subsequently ligated into *Sna*BI-digested and dephosphorylated pG8SAET vector.

Approximately 4.5 x 10^6 ampicillin-resistant transformants were obtained after electrotransformation of the ligated material into *E. coli*, TG1 cells.

Twenty randomly picked transformants were all shown to contain inserts. Cells from an overnight culture of the transformants were infected with helper phage R408 and poured together with soft agar onto LA + ampicillin plates (LB medium supplemented with 1.5 % agar and 50 µg ampicillin/ml) and incubated overnight. Phage particles were eluted from the soft agar by addition of LB and vigorous shaking. The suspension was centrifugated and the supernatant sterile filtrated. The titer of the library was determined to 7 x 10^{10} CFU/ml.

This library was used in the following Section B, wherein phage particles containing inserts related to Fn-binding properties are identified.

B. Panning of the phagemid library. Microtiter wells (Maxisorp, Nunc, Copenhagen, Denmark) were coated with human Fn (Sigma, St. Louis, MO) at a concentration of 100 µg/ml in 50 mM sodium carbonate, pH 9.7. The wells were blocked with PBS-0.05% Tween 20 (PBS-T) containing casein (0.1mg/ml). After washing the wells with PBS-T, the library from Section A above was added to the wells. Thereafter, the wells were extensively washed with PBS-T and then eluted with 140 mM NaCl, 50 mM Na-citrate (pH 2.0).

To obtain transformed cells containing the Fn-binding insert, eluate was collected, neutralized, infected with *E .coli* TG1 cells and spread on LA plates containing ampicillin. Next day, approximately 1,500 colonies were pooled and after infection with helper phage R408 the sample was mixed with soft agar and poured out on LA plates. After incubation overnight, the phagemid particles were extracted and subjected to another round of panning. After the first panning, 41 % of the colonies were found to bind Fn and after the second panning all 180 colonies were positive for Fn-binding.

C. Screening for Fn-binding clones containing an insert encoding SFS. In this section, the gene encoding the novel Fn-binding protein SFS was screened for based on the knowledge that the previously disclosed *fnz* gene isolated from subspecies *zooepidemicus* and encoding the Fn-binding protein FNZ is present also in the genome of subspecies *equi* . Thus, a negative screening test was used, wherein cells containing inserts related to Fn-binding, i. e. cells positive when screening with a rabbit anti-Fn antibody, but not containing inserts related to Fn-binding of FNZ, i. e. those of the positive cells that were negative when the *fnz* gene is used as a screening probe, were selected and presumed to contain SFS-related inserts.

From each panning performed in section B, 180 colonies were transferred in triplicate to LA plates and incubated overnight. The following day, one plate was stored

(masterplate) and the colonies from the two remaining plates were transferred to nitrocellulosa filters and incubated for two hours.

Cells from one filter were lysed using chloroform vapor and after blocking the filter with PBS-T supplemented with casein (0.1mg/ml), it was incubated with human Fn (1μg/ml; Sigma) for 2 h, and after washing, a rabbit anti-Fn antibody (diluted 1/1000; Sigma) was added. After 1 h of incubation and washing, the filter was incubated for additional 1 h with a HRP-labeled secondary antibody (diluted 1/1000; Bio-Rad, Richmond, Calif.). Reactive bands were visualized by using 4-chloro-1-naphthol (Serva, Heidelberg, Germany).

The second filter was subjected to colony hybridization essentially as described in Sambrook et al. (22) with use of a radioactively labeled probe that covered the entire *fnz* gene and was generated by PCR amplification of chromosomal DNA from subsp. *zooepidemicus* strain ZV using the primers:

5-fnz, 5′-CGGGATCCCTATTACACATTCTCATCTCATAT (positions 19-42) and

3-fnz, 5′-GGAATTCCAGAAAGCCCGCCTGTAAAC (positions 1954-1935).

The indicated positions in the respective primers correspond to the published sequence of the gene *fnz* (9).

In these colony hybridization tests, 41 % of Fn-binding clones from the first panning and 30 % from the second panning hybridized against the *fnz* gene from subsp. *zooepidemicus* ZV used as a probe.

D. Cloning and isolation of the gene *sfs*. Clones from Section C, displaying Fn-binding activity but negative in the colony hybridization assay, were selected as candidate *sfs*-gene-containing clones. Accordingly, these clones were sequenced using thermo sequence dye terminator cycle sequencing pre-mix kit (Amersham) and the ABI Model 377XL DNA sequencer. Computer programs from the PCGENE, DNA, and protein sequence analysis software package (Intelligenetics, Inc., Mountain View, Calif.) were used to record and analyze the sequence data.

Altogether, eleven Fn-binding but *fnz* negative clones were analyzed and found to contain inserts identical to one of four different types of inserts, all with overlapping sequences and an open reading frame. These four different phagemid clones designated S1-S4 are shown in Fig 1B. Based on the overlapping sequences, primers were designed and used to generate a probe consisting of the complete *sfs* gene using PCR amplification as disclosed in the section concerned with Southern blots below.

To isolate the complete gene encoding the Fn-binding activity, Southern blot analysis of restriction enzyme digested chromosomal DNA of subsp. *equi* Bd 3221 was performed as

disclosed below and revealed that a 2.6 kb *Ssp*I fragment contained *sfs*. Accordingly, fragments of this size were purified from a preparative agarose gel and ligated into pUC19. The ligation mix was electroporated into *E. coli* and transformants were screened for Fn-binding activity as described above. Among several positive clones one, designated pSFS62, was selected and the insert sequenced.

This clone, pSFS62, had an open reading frame of 1,035 bp, from which the phagemid sequences were found to originate (Figs. 1 and SEQ. ID. NO. 2). The open reading frame is preceded by sequences typical for promoter and ribosome-binding sites (not shown) and is followed by sequences (not shown) resembling a transcriptional termination, suggesting that the gene is translated from a monocistronic messenger. The SFS-coding nucleotide sequence of the *sfs* gene is shown in SEQ. ID. NO. 2.

Southern blots. The Southern blot analysis referred to above and further below, was performed according to the following. Agarose imbedded chromosomal DNA digested with *Apa*I was resolved on 1.2% SeaKem GTG agarose gel (FMC, Rockland, ME) in 0.5 x TBE buffer by PFGE using a Gene Navigator (Pharmacia Biotech, Uppsala, Sweden) as earlier described (10). The DNA was transferred to nylon filters (Hybond-N+, Amersham) by vacuum blotting (VacuGene XL, Pharmacia Biotech) in accordance with the manufacturer's protocol. After cross-linking, the filters were prehybridized for 2 h at 65°C in 6 x SSC, 3 x Denhardt's solution, and 0.5% SDS and subsequently incubated with the radioactively labeled *sfs* probe overnight, using the same conditions. The membranes were washed 3 x 20 min at 65°C with 0.2 x SSC, 0.1% SDS and subjected to autoradiography.

The probe *sfs* was generated by PCR amplification of chromosomal DNA from subsp. *equi* Bd 3221 using the primers:

fs5, 5'-ACAAGCCATGGAGCACTTGTCTTTGGAGGT and

fr4, 5'-GTCGGGATTGTAAGAATAGCC.

The single band obtained after agarose gel electrophoresis was purified and random-primed.

Example 2. Construction and purification of SFS as a fusion protein. The purified PCR-fragment from Example 1 encoding the mature protein of SFS, and described under Southern blots in Example 1, Section D, was digested with *Nco*I and ligated into *Sna*BI-*Nco*I opened pG8SAET. This vector encodes a 13 amino acid peptide tag (E-tag) which facilitates the purification of the recombinant protein using a HiTrap Anti-E tag column (Pharmacia Biotech). The recombinant protein SFS-E was purified from the periplasmic space according to the manufacturer's protocol.

After cleaving from the E-tag, the SFS protein was obtained having a calculated molecular mass of 40 kDa. The charged amino acids, followed by a stretch of hydrophobic residues in the N-terminal end of the protein, indicate a signal sequence and by the method of von Heijne (14) a possible signal sequence cleavage site was found between amino acids 29 and 30, resulting in a mature protein with a calculated molecular mass of 36 kDa. The isolated Fn-binding phagemid clones contained inserts originating from the central part of the protein, where two repetitive sequences of 21 residues, called R1 and R2, resp., are situated (Fig. 1). Three amino acids were found to dominate the composition of protein SFS, 53 residues are glycines (14.4%), 39 serines (10.6%), and 38 prolines (10.3%). These three amino acids are evenly distributed in the protein in contrast to the 13 tyrosine residues which occur only in the C-terminal part of the protein. Protein SFS does not contain any sequence motifs known to mediate attachment to the bacterial cell-wall.

Example 3. Inhibition assays. Cells from overnight cultures of streptococci were collected by centrifugation, washed in PBS, and suspended in PBS-0.2% Tween 20 to an optical density at 600 nm of 0.2. In cases of inhibition, 25 nM of affinity purified fusion protein SFS-E was preincubated 15 min with 16 pM of ^{125}I-labeled human Fn (91,061 cpm) and thereafter bacteria (500 µl) were added. After two hours incubation at room temperature, the mixtures were centrifuged and the supernatants removed. The radioactivity associated with the pellets was quantified in a gamma counter (LKB Wallac, Turku, Finland). Radioactivity (808 cpm) recovered from a control (tubes that contained no streptococci) was subtracted from each test.

Example 4. Expression of the *sfs* gene. RNA was extracted from *S. equi* cells by using the Blue FastRNA kit (Bio 101, Vista, CA) according to the manufacturer's protocol. RNA concentration was determined spectrophotometrically and by visual estimation of the rRNA bands on an agarose gel. RNA (10 µg) was loaded on a formaldehyde-containing agarose gel. RNA was transferred by vacuum-blotting to a positively charged nylon filter (Hybond-N+, Amersham) and cross-linked. Further steps were performed as described for Southern blots above with the exception that ssDNA was added to the pre-hybridization and hybridization solutions.

Example 5. The ability of SFS to inhibit the binding between collagen and Fn. For the enzyme-linked immunosorbent assay (ELISA), polystyrene 96-well microtiter plates were coated for one hour with collagen type I from calf skin (Boehringer, Mannheim, Germany) in PBS. The wells were blocked for one hour with PBS-T supplemented with casein (0.1 mg/ml) and then washed four times with PBS-T. The fusion protein SFS-E was

diluted in a two-fold serial and added to the wells together with 0.2 ng of Fn. After 2 h incubation, the wells were washed and a rabbit anti-Fn antibody was added and allowed to bind for 1 h. Finally, the wells were incubated for 1 h with a secondary HRP-labeled antibody. After washing, bound material was quantified by using tetramethylbenzidine (Boehringer) and a microplate reader (Bio-Tek Instruments, Vinooski, VT). Measurement was done at a wavelength of 450 nm. Absorbancy in wells without added fusion protein was set to 100 % and absorbancy in wells where Fn had been excluded was set to 0 %.

RESULTS

I. The gene *sfs* is generally present in isolates of subsp. *equi*. Southern blots performed as disclosed above revealed that a [^{32}P] dATP-labeled probe, corresponding to the gene *sfs*, hybridized to all the 50 subsp. *equi* and to 41 out of 48 subsp. *zooepidemicus* isolates tested. The results from the hybridization analysis are, for a selected number of strains, shown in Fig. 2. No significantly weak signal, that could not be explained by less chromosomal DNA on the gel, was detected for any of the positive *S. equi* isolates. The seven isolates of subsp. *zooepidemicus* that were *sfs* negative could not be related to each other, considering symptoms, temporal, and geographical origin. Furthermore, the seven negative isolates were obtained from different species, horses (n = 4), cows (n = 2), and dog (n = 1). The *sfs* probe did not hybridize to any of the three control strains of other streptococcal species (Fig. 2).

II. Protein SFS displays sequence similarity to both collagen and a potential cell-wall protein of *S. pyogenes*. Collagen sequences gave highest scores when searching the database Swissprot for SFS-like sequences. The similarity was evenly distributed through protein SFS, and the main reason for the high score is the high content of glycine, serine, and proline, i. e. residues which are also common in collagen. However, a more pronounced similarity was seen for the Fn-binding domain of SFS against collagen. A sequence comparison was also done against the Oklahoma *S. pyogenes* genomic sequence database, which at the time of search consisted of 98% of the *S. pyogenes* genome. SFS aligned best against a database sequence which besides high content of glycine and proline residues also displayed the motif QGERGETGP. Eight of these nine residues are present in the Fn-binding domain of SFS. Similar motifs are also present in chains of collagen. Alignment of these sequences are shown in the following Table I. At a closer study of the aligned *S. pyogenes* sequence, it was found that the aligned motif is situated in the middle of a potential gene, encoding a typical streptococcal cell-surface protein. This statement is based on the following: (i) promoter sequences and a putative ribosome-binding site are present adjacent to an open reading frame, (ii) in the C-terminal part there is a proline-rich domain with the cell-wall

anchoring motif LPXTGX, (iii) the LPXTGX motif is directly followed by a stretch of 23 hydrophobic residues and the open reading frame is terminated by six residues whereof three are charged, and (iv) a potential hairpin loop is situated 38 bp downstream the stop codon. However, a start codon in an acceptable distance to the ribosome binding site could not be found.

Table I

SFS	**QGERGEAGPP**
S. pyogenes	**QGERGETGPA**
Collagen α2 (I) 711	**PGERGEVGPA**
Collagen α1 (I) 991	**SGERGPPGPM**
Collagen α1 (II) 329	**PGERGRTGPA**
Collagen α1 (III) 797	**PGERGETGPP**
Collagen α1 (IV) 319	**QGEKGEAGPP**

In this table, alignment of amino acid sequences from different types of collagen and the potential cell-surface protein from *S. pyogenes* to a motif present in the Fn-binding domain of SFS is illustrated. The figures indicate the number of the first amino acid from the collagen sequences. Bold letters indicate identical residue to the SFS motif.

III. Inhibition of binding between Fn and cells of *S. equi*. Recombinant protein SFS was purified by using affinity tails and the purified protein was found to bind Fn in a Western blot assay (data not shown). Before adding iodinated Fn to cells of *S. equi* the labeled Fn was, in appropriate cases, preincubated with SFS in a molar ratio of 1:1,500. After incubation the cells were collected by centrifugation and after removing the supernatant, the radioactivity bound to the pellets was measured. From the results from the inhibition experiments shown in Fig. 3 it is evident that the protein SFS has, for both subspecies, an inhibitory effect, although the two subsp. *equi* strains bind considerable less Fn compared to the two subsp. *zooepidemicus* strains

IV. Protein SFS inhibits the binding between Fn and collagen. The similarity between protein SFS and collagen suggested that these proteins might bind to the same site on the Fn molecule. In order to investigate this, microtiter wells coated with collagen were incubated with a mixture of Fn and a serial dilution of protein SFS. Bound Fn was detected by an anti-Fn antibody and as seen in Fig. 4, protein SFS inhibits the binding in a concentration dependent way. In a similar assay, the previously known protein FNZ did not inhibit the binding between Fn and collagen (data not shown). Furthermore, protein SFS did not inhibit

the binding between the said protein FNZ and Fn, and the protein FNZ did not inhibit the binding between protein SFS and Fn. Protein SFS does not bind collagen. This was tested in order to control that the inhibition of binding between Fn and collagen by protein SFS is dependent on the binding of protein SFS to Fn and not to collagen. Taken together this suggests that protein SFS and the previously known protein FNZ have clearly separate binding sites on the Fn molecule and that protein SFS binds to the 30-40 kDa collagen-binding domain of Fn.

In the following Example 6, the potential use of the present protein as a vaccine is illustrated in a test wherein the immunogenic properties of the present novel protein are confirmed.

Example 6. Immunogenic properties of Protein SFS. Affinity purified recombinant protein SFS (Example 2) was, under reducing conditions, subjected to SDS-PAGE on a precasted 8-25% gradient-gel using the PHAST system (Pharmacia Biotech, Sweden). The molecular weight markers used were obtained from BioRad, CA, USA. After electrophoresis was completed, a nitrocellulose (NC) filter (Hybond C, Amersham, UK) previously soaked in PBS was put on the gel and the temperature raised to 45° C. After 45 minutes, the NC-filter was wetted with 1 ml PBS, and removed and placed in 15 ml PBS-T containing casein (0.1 mg/ml) for 1 hour (with two changes of PBS-T casein solution) at room temperature under gentle agitation.

The gradient-gel was after transfer removed and stained with Coomassie-blue using the PHAST system. The NC-filter was removed and incubated in 5 ml PBS-T casein solution containing 5 µl serum from a horse which previously had got the diagnosis strangles and found to be a carrier of *S. equi*. After 2 hours incubation at room temperature, under gentle agitation, the filter was extensively washed with PBS-T and incubated in 5 ml PBS-T casein solution containing rabbit anti horse antibodies (Nordic Immunology, Netherlands) at a dilution of 1:1000. After 1 hour of incubation at room temperature, under gentle agitation, the filter was extensively washed with PBS-T and incubated in 5 ml PBS-T casein solution containing horseradish peroxidase labeled goat anti rabbit IgG (BioRad) at a dilution of 1:1000. After 1 hour of incubation at room temperature, under gentle agitation, the filter was extensively washed with PBS-T and PBS.

To visualize the bound IgG conjugate, the filter was transferred to a solution containing a substrate for peroxidase (containing 25 ml PBS + 6ml 4-chloro-1-naphtol (Sigma, USA, 3 mg/ml in methanol) + 20 µl H_2O_2 (35%). After about 15 minutes, the

degree of color was measured by eye. The bands appearing on the NC-filter and the bands appearing on the corresponding Coomassie-blue stained PAGE were compared.

The obtained results clearly showed that (i) the recombinant produced SFS protein was recognized by antibodies present in the serum from the horse with strangles, and (ii) no bands were seen on the NC-filter in the lane with the different molecular weight markers. Thus, this means that protein SFS is expressed by *S. equi* during the infection process and that this protein is immunogenic

REFERENCES

1. **Cue, D., P. E. Dombek, H. Lam, and P. P. Cleary.** 1998. *Streptococcus pyogenes* serotype M1 encodes multiple pathways for entry into human epithelial cells. Infect. Immun. **66**:4593-4601.
2. **Galán, J. E., and J. F. Timoney.** 1988. Immunologic and genetic comparison of *streptococcus equi* isolates from the united states and europe. J. Clin. Microbiol. **26**:1142-1146.
3. **Hanski, E., and M. G. Caparon.** 1992. Protein F, a fibronectin-binding protein, is an adhesin of the group A streptococcus *Streptococcus pyogenes*. Proc. Natl. Acad. Sci. USA **89**:6172-6176.
4. **Hanski, E., P. A. Horwitz, and M. G. Caparon.** 1992. Expression of protein F, the Fibronectin-binding protein of *Streptococcus pyogenes* JRS4, in heterologous streptococcal and enterococcal strains promotes their adherence to respiratory epithelial cells. Infect. Immun. **60**:5119-5125.
5. **Hynes, R. O.** 1990. Fibronectins. Springer Verlag, New York.
6. **Jacobsson, K., and L. Frykberg.** 1996. Phage display shot-gun cloning of ligand-binding domains of procaryotic receptors approaches 100% correct clones. BioTechniques **20**:1070-1081.
7. **Jadoun, J., V. Ozeri, E. Burstein, E. Skutelsky, E. Hanski, and S. Sela.** 1998. Protein F1 is required for efficient entry of *Streptococcus pyogenes* into epithelial cells. J. Infect. Dis. **178**:147-158
8. **Jorm, L. R., D. N. Love, G. D. Bailey, G. M. Bailey, and D. A. Briscoe.** 1994. Genetic structure of populations of β-haemolytic Lancefield group C streptococci from horses and their association with disease. Res. Vet. Sci. **57**:292-299.
9. **Lindmark, H., K. Jacobsson, L. Frykberg, and B. Guss.** 1996. Fibronectin-binding protein of *Streptococcus equi* subsp. *zooepidemicus*. Infect. Immun.

64:3993-3999.

10. **Lindmark, H., P. Jonsson, E. Olsson Engvall, and B. Guss.** 1998. Pulsed-field gel electrophoresis and distribution of the genes *zag* and *fnz* in isolates of *streptococcus equi*. In Press.
11. **Molinari, G., S. R. Talay, P. Valentin-Weigand, M. Rohde, and G. S. Chhatwal.** 1997. The fibronectin-binding protein of *Streptococcus pyogenes* SfbI, is involved in the internalization of group A streptococci by epithelial cells.Infect. Immun. 65:1357-1363.
12. **Nilsson, M., L. Frykberg, J. I. Flock, L. Pei, M. Lindberg, and B. Guss.** 1998. A fibrinogen-binding protein of *Staphylococcus epidermidis*. Infect. Immun. 66:2666-2673.
13. **Skjold, S. A., P. G. Quie, L. A. Fries, M. Barnham, and P. P. Cleary.** 1987. DNA fingerprinting of *streptococcus zooepidemicus* (Lancefield group C) as an aid to epidemiological study. J. Infect. Dis. **155**:1145-1150.
14. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14:4683-4690.

Nucleotide Sequence Accession Number: The nucleotide sequence of the *fnz* gene (9) is available from the EMBL sequence data bank under accession number X99995. The complete gene sequence and the deduced amino acid sequence of the protein FNZ are shown below:

SUBSTITUTE SHEET (RULE 26)

→ A2 (above line 1200), → R1

CTATAGTACAACAAAACTATGGTAGTACCGAAGGATATCATGGTCAATCAGGCATTCTTGAGCAAACCCAACATACTAACCCACCTGGTATCATACTAGGCGGCTCACCGAAATCTCCAAA 1200
P I V E Q N Y G S T E G Y H G Q S C I L E P T P D T N P P G I I L C G S G N V S 364

→ R2

CGCATGAACATACTACAAACCCTCATCTCATCCCCATCGGCGGCGGTCTAGCTGGCGAATCAGGAGAAACGACACCTAAACCACCACAAACCGGCGGCGGCAAGGACCAGTCATCGAGACAA 1320
T H E D T R N D N L N G I G G G L A G E S C E T T P K P G Q T G G Q C P V I E T 404

→ R3

CAGAGGATACACAAAAAGGCATGTCTGCACAATCTCCTCCCACTATCCCATCTCAGAAAACACCAAAAAGCCGGAGGTCATGATTGGTGGTCAGCCACAAACCATCGAGACAACAGAGGACA 1440
T E D T Q K C N S C Q S G G T I S S E N T K K P E V N I G G Q G Q T I E T T E D 444

→ R4

CACAAAAAGGCATGTCTGGACAATCTGGCGGTACTATCGAGTCAGAGCACACTAACAAACCTCACGTCATGATTGGTGGTCAGGGACAAATCATCGACTTCTCTGAAAACACCCAATCAG 1560
T Q K G N S G G G S C C T I E S E D T K K P E V N I G G Q G Q I I D P S E N T Q S 484

→ R5

GTATGTCTCGCCAGTCTGGTCACACTACGGTAATTGAGGATACCAAGAAGTCTGACATAATCATTGGTGGCCAAGGACAAATCATCGACTTCTCTGAGGATACTCAGCCCGGTATGTCTC 1680
G M S G Q S G D T T V I E D T K K S E I I I G G Q G Q I I D P S E D T Q P G M S 524

← W

GTCAATCTGGACGGCACTACAATTGTCGAAGACACCAAGAAGCCGACACCTAAGCCTAAACCTGCACCTGCGCCAATTCTTAATCACCAAAAAACGTAACAAAGGCACTCATCTCCCACAGA 1800
C Q S G G T T I V E D T K K P T P K P K P A P A P I V N D E K P N K C T H L P Q 564

→ M

CAAGTGATATGAAGCAACTCACCCTAAGCATCATCCCTGCAATGTCAATGCTGCTTGTCCTATGTCTGTCTCTATTCAAGCGACCATCTAAAAAACACTAAGCAACACTGACAAGTGATC 1920
T E D M K Q L T L S I I G A M S M L L V L C L S L P K R D S K K D * 597

ATTATTGTACGGACGTTTACAGGCGGGCTTTCTCCCCTCTTAAAACACCCTACAACAGCAAAACGGTCTAAATAGAAATATGCCTCAAGGCTCAGACAATGGCTCTAGCCCCTGGGGCCT

TTTTTCACATTTATAATACAAGAGAGCAAGCAGTGACACTTTCTTTTGCCCTATATCTCTTAACACTAGCCCTATGTGCAGTTAC

CLAIMS

1. A protein having an amino acid sequence encoded by a nucleic acid sequence , that can be isolated from and forms a portion of the genomes of *S. equi*, and which protein can be expressed from said nucleic acid sequence and binds specifically to mammal fibronectin, or an analog or fragment of said protein.

2. An isolated protein specifically binding to fibronectin and having an amino acid sequence as shown in SEQ. ID. NO. 1, or an analog or fragment thereof.

3. The protein of claim 2 having an amino acid sequence as shown in SEQ. ID. NO. 1 containing deletions or substitutions of amino acids.

4. The protein of claim 2, wherein the protein is a fragment comprised of the amino acid sequence of SEQ. ID. NO. 1 that lacks an N-terminal sequence, suitably amino acids no 1-19 inclusive in the sequence of SEQ. ID. NO. 1.

5. The protein of claim 3, which has an amino acid sequence corresponding to a portion of the sequence as shown in SEQ. ID. NO. 1 wherein an amino acid sequence binding to a collagen-binding domain of fibronectin (Fn) and comprising the sequence consisting of amino acids QGERGEAGPP, is deleted.

6. The protein of claim 1, wherein the protein has an amino acid composition of approximately 53 glycine residues, 39 serine residues and 38 proline residues evenly distributed in the protein and optionally 13 tyrosine residues in the C-terminal part of the protein.

7. A DNA fragment comprising a nucleotide sequence coding for a protein according to any one of claims 1-6.

8. The DNA fragment of claim 7, wherein said fragment has a nucleotide sequence as shown in SEQ, ID. NO. 2 or an equivalent thereof.

9. A recombinant DNA molecule comprising a replicable vector, which suitably is an expression vector, and a DNA fragment according to claim 7 or 8 inserted therein.

10. A host cell comprising a DNA fragment in accordance with claim 7 or 8,or the recombinant DNA molecule of claim 9.

11. The host cell of claim 10, wherein said cell is a prokaryotic host cell, suitably a prokaryotic host cell comprised of a strain of *E. coli*.

12. A method of producing the protein of any one of claims 1-6, or fragments or analogs thereof comprising culturing a host cell as defined in claim 10 or 11, and isolating the expression product comprising the protein from the culture.

13. The method of claim 12, wherein said method further comprises purification of the expression product, such as by affinity chromatography.

14. The method of claim 12, which method comprises

(a) introducing the DNA fragment encoding the protein or fragment or analog thereof into an expression vector;

(b) introducing the said vector, which contains the said DNA fragment, into a compatible host cell;

(c) culturing the host cell provided in step (b) under conditions required for expression of the product encoded by said DNA fragment; and

(d) isolating the expressed product from the cultured host cell, and, optionally,

(e) purifying the isolated product from step (d) by affinity chromatography or other chromatographic methods known in the art.

15. A vaccine comprising the fibronectin-binding protein or a fragment or an analog thereof as defined in any one of claims 1-6 or as produced by a method as defined in any one of claims 12-14.

16. The vaccine of claim 15, which vaccine is a vaccine that protects horses against strangles caused by *S. equi* infection.

17. Use of a fibronectin-binding protein, analogs or fragments thereof in the preparation of a vaccine protecting against S. *equi* infection inclusive of strangles caused by subsp. *equi* infection in horses.

18. An antibody specific for a fibronectin-binding protein of any one of claims 1-6 or a fragment or an analog thereof, which antibody is polyclonal or monoclonal, or a fragment of said antibody.

19. An antigenic preparation comprising an antigen consisting of the protein of any one of claims 1-6 or a fragment or an analog thereof.

20. An antiserum comprising an antibody of claim 18, which is comprised of a polyclonal antibody.

21. A method for the production of an antiserum, said method comprising administering an antigenic preparation of claim 19 to an animal host to produce antibodies in said animal host and recovering antiserum containing said antibodies produced in said host animal.

22. A method of prophylactic or therapeutic treatment of *S. equi* infection in mammals, suitably horses, comprising administering an immunologically effective amount of a vaccine of claim 15 or 16, an antibody of claim 18, or an antiserum of claim 20.

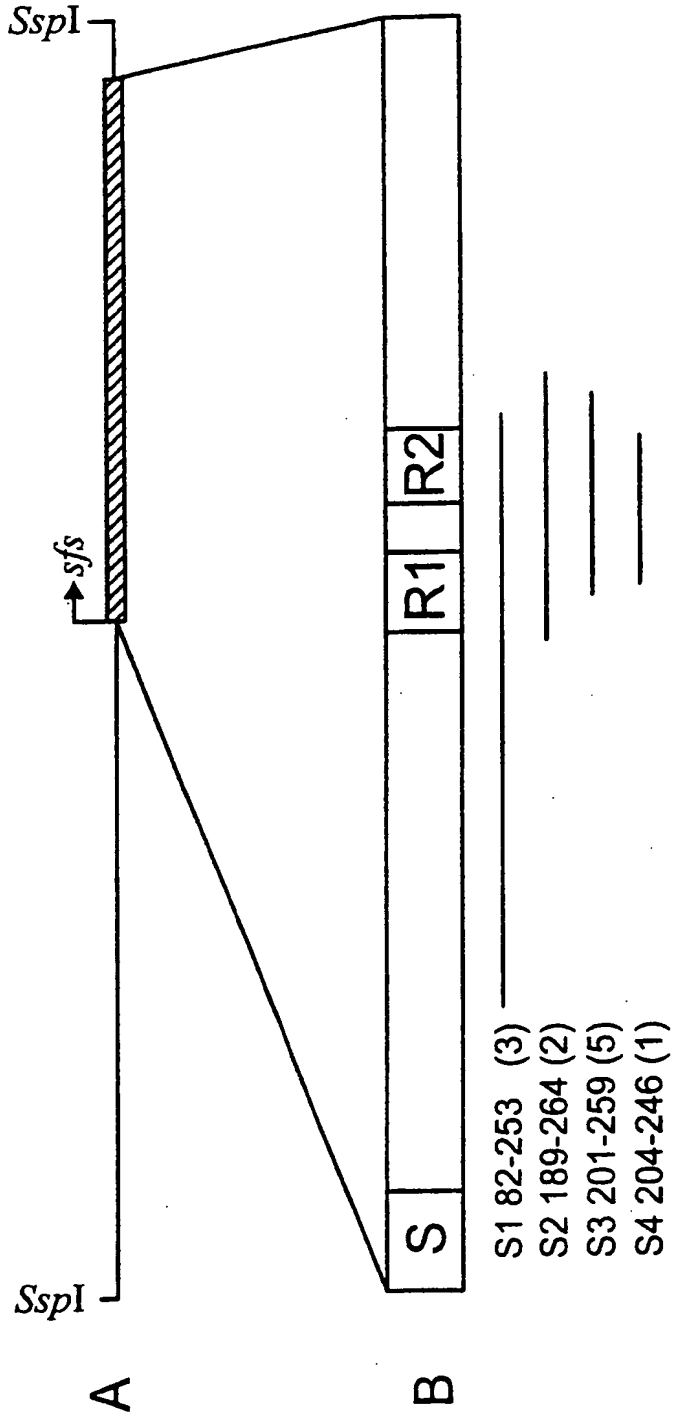
SspI
sfs
SspI
R2
R1
S
S1 82-253 (3)
S2 189-264 (2)
S3 201-259 (5)
S4 204-246 (1)
A
B

FIG.1

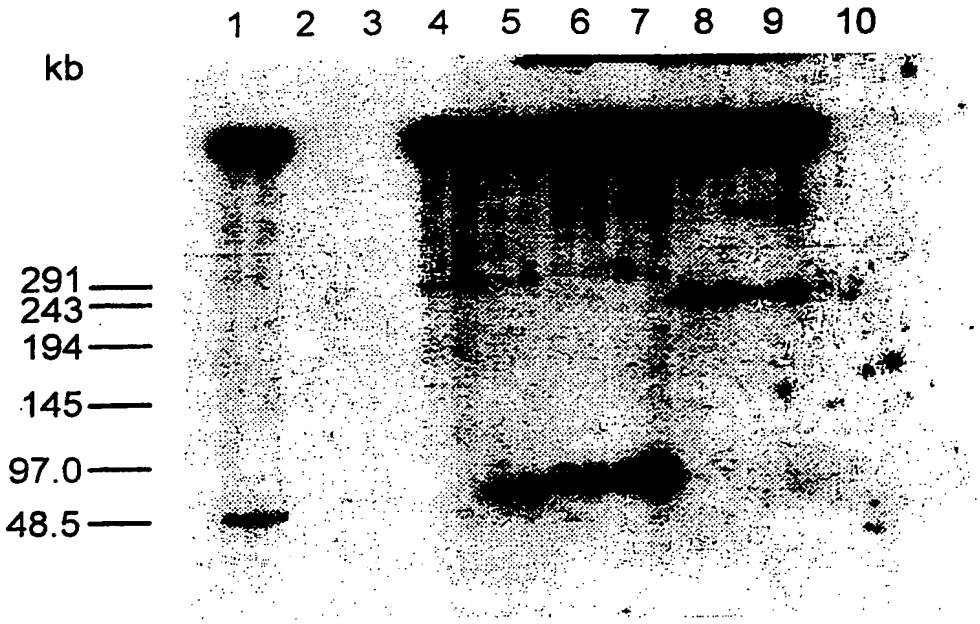

FIG.2

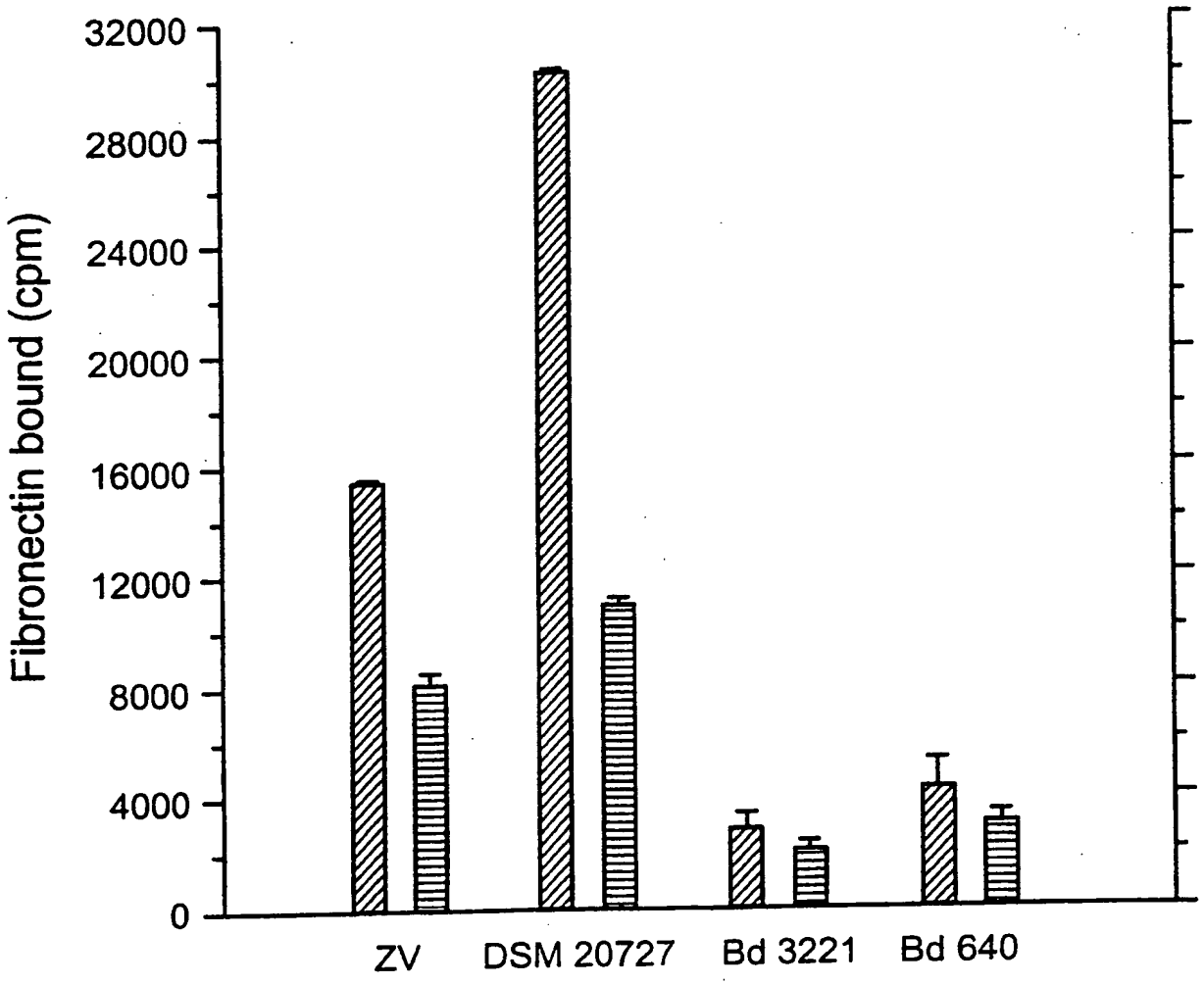
Fibronectin bound (cpm)
32000
28000
24000
20000
16000
12000
8000
4000
0
ZV
DSM 20727
Bd 3221
Bd 640

FIG.3

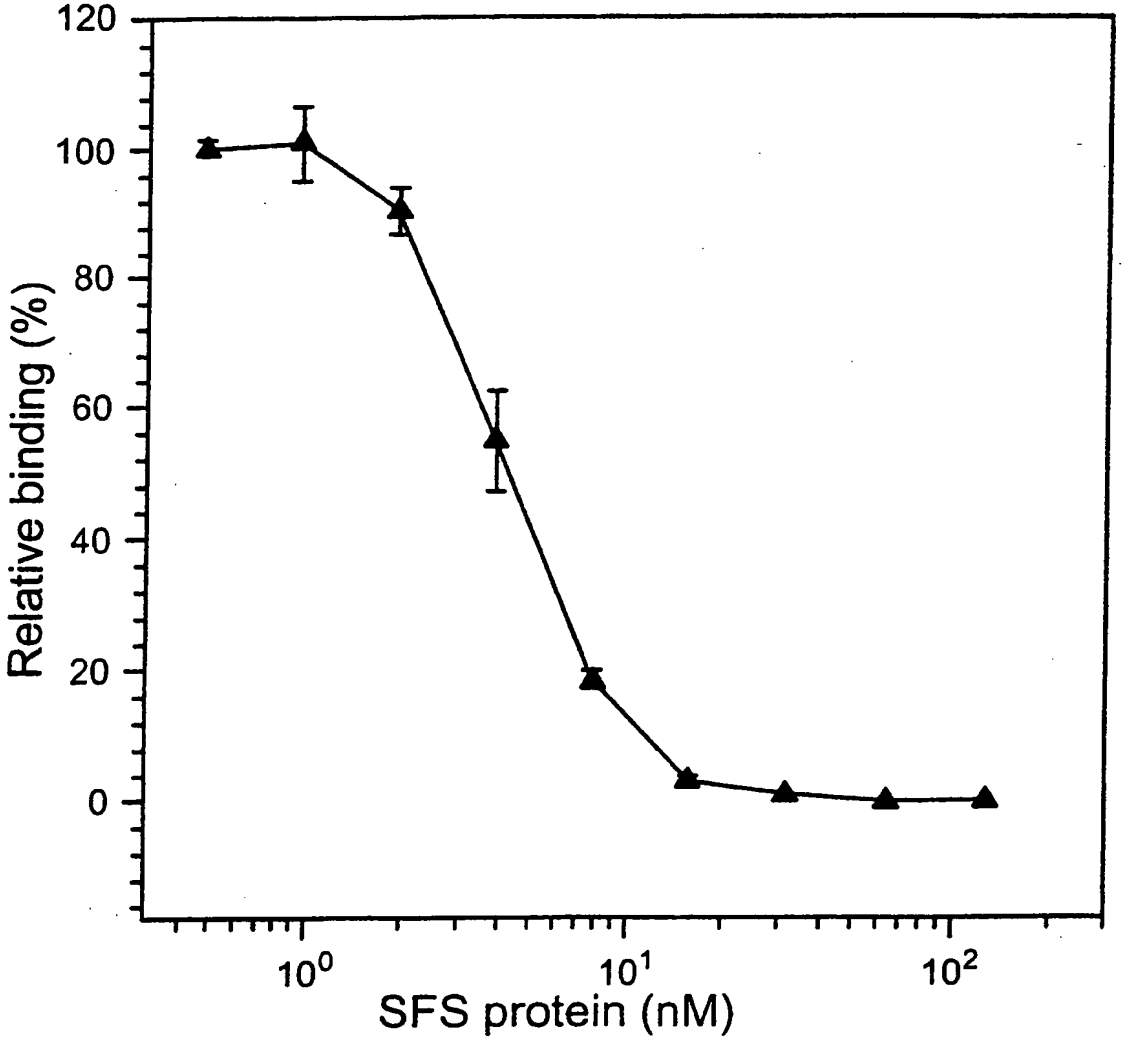
120
100
80
60
40
20
0
Relative binding (%)
10^0
10^1
10^2
SFS protein (nM)

FIG.4

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Bengt Guss et al
(B) STREET: Dag Hammarskjolds vag 238 B
(C) CITY: Uppsala
(E) COUNTRY: Sweden
(F) POSTAL CODE (ZIP): SE-756 52

(ii) TITLE OF INVENTION: Novel fibronectin-binding protein.

(iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 281 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Streptococcus equi

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
Met Arg Lys Thr Glu Gly Arg Phe Arg Thr Trp Lys Ser Lys Lys Gln
1               5                  10                  15

Trp Leu Phe Ala Gly Ala Val Val Thr Ser Leu Leu Leu Gly Ala Ala
            20                  25                  30

Leu Val Phe Gly Gly Leu Leu Gly Ser Leu Gly Gly Ser Ser His Gln
        35                  40                  45

Ala Arg Pro Lys Glu Gln Pro Val Ser Ser Ile Gly Asp Asp Asp Lys
    50                  55                  60

Ser His Lys Ser Ser Ser Asp Gln Pro Thr Asn His Gln His Gln Ala
65                  70                  75                  80

Thr Ser Pro Ser Gln Pro Thr Ala Lys Ser Ser Gly His His Gly Asn
                85                  90                  95

Gln Pro Gln Ser Leu Ser Val Asn Ser Gln Gly Asn Ser Ser Gly Gln
            100                 105                 110
```

Ala Ser Glu Pro Gln Ala Ile Pro Asn Gln His His Gln Pro Gln Gly
115 120 125

Lys Pro Gln His Leu Asp Leu Gly Lys Asp Asn Ser Ser Pro Gln Pro
130 135 140

Gln Pro Lys Pro Gln Gly Asn Ser Pro Lys Leu Pro Glu Lys Gly Leu
145 150 155 160

Asn Gly Glu Asn Gln Lys Glu Pro Glu Gln Gly Glu Arg Gly Leu Pro
165 170 175

Gly Leu Asn Gly Glu Asn Gln Lys Glu Pro Glu Gln Gly Glu Arg Gly
180 185 190

Glu Ala Gly Pro Pro Ser Thr Pro Asn Leu Glu Gly Asn Asn Arg Lys
195 200 205

Asn Pro Leu Lys Gly Leu Asp Gly Glu Asn Lys Pro Lys Glu Asp Leu
210 215 220

Asp Gly Tyr Asn His Gly Arg Arg Asp Gly Tyr Arg Val Gly Tyr Glu
225 230 235 240

Asp Gly Tyr Gly Gly Lys Lys His Lys Gly Asp Tyr Pro Lys Arg Phe
245 250 255

Asp Glu Ser Ser Pro Lys Glu Tyr Asn Asp Tyr Ser Gln Gly Tyr Asn
260 265 270

Asp Asn Tyr Gly Asn Gly Asn Pro Asp
275 280

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 846 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Streptococcus equi

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATGAGAAAAA CAGAAGGACG TTTTCGCACA TGGAAGTCCA AAAAACAATG GCTATTTGCC 60

GGTGCAGTGG GAGCTGCACT TGTCTTTGGA GGTTTATTAG GAAGTCTTGG TGGCTCATCC 120

CAGCAGCCAG TCAGCTCGAT TGGAGATGAC GATAAGTCGC ACAAGAGCTC ATCACCACCG 180

AAAAAGGATA ACTTGCAGCC TAAGCCTTCA GATCAGCCTA CTAATCGCCC GTCCCAGCCG 240

ACAGCAAAGA GCTCAGGTCA TCATGGGAAT CAACCTCACC AAGGAAATAG TAGTGGACAG 300

GCCTCAGAGC CTCAGGCTAT TCCTAATCAA GGGCTGAGAG GAGGTAACAG CTCTGGTTCA 360

GGTCATCACC ATCAGCCACA AGATCTAGGT AAGGATAATT CTAGCCCGCA GCCTCAACCA 420

AAGCCTCAGG GCAAAAAAGG CTTGAATGGT GAAAATCAGA AGGAACCGGA GCAAGGTGAA 480

CGAGGTTCAG GGTTGAGTGG TAATAATCAA GGCCGTCCTT CGCTTCCAGG CTTGAATGCA 540

GAGCAAGGTG AACGAGGTGA AGCCGGTCCC CCATCAACTC CGAATTTAGA TCCTTTAAAA 600

GGATTAGATG GAGAGAATAA GCCAAAGGAA GATTTAGACG GTAATGATGA ATCACCAAAA 660

CTTAAAGACG AACACCCCTA CAATCATGGT CGTCGCTATG AGGATGGATA TGGTGGCAAA 720

AAGCACAAAG GAGATTATCC TAAGCGAAAG GAATATAATG ACTATAGTCA AGGGTATAAT 780

GATAATTATG GAAATGGCGA TAGAGGTGGT AAGAGAGGAT ACGGCTATTC TTACAATCCC 840

GACTAA 846

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 99/02448

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: C07K 14/315, A61K 39/09, C07K 16/12
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: C07K, A61K, C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Infection and Immunity, Volume 64, No 10, 1996, Hans Lindmark et al, "Fibronectin-Binding Protein of Streptococcus equi subsp. zooepidemicus" page 3993 - page 3999	1,6-22
A	--	2-5
A	WO 9831389 A2 (THE TEXAS A & M UNIVERSITY SYSTEM), 23 July 1998 (23.07.98) --	1-22
P,X	Infection and Immunity, Volume 67, No 5, May 1999, Hans Lindmark et al, "SFS, a Novel Fibronectin-Binding Protein from Streptococcus equi, Inhibits the Binding between Fibronectin and Collagen" page 2383 - page 2388 --	1-22

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" erlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 April 2000	08 -05- 2000
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Patrick Andersson/EÖ Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE99/02448

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 2
because they relate to subject matter not required to be searched by this Authority, namely:
See PCT Rule 67.1.(iv).: Methods for treatment of the human or animal body by surgery or therapy, as well as diagnostic methods.

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

02/12/99

International application No.
PCT/SE 99/02448

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9031389 A2	23/07/98	NONE	